

# Stock Market Analysis and Prediction

Pranit Brahmhatt<sup>1</sup>, Kushang Patel<sup>2</sup>

<sup>1</sup>Student, Dept. of Computer Engineering, SVIT, Gujarat, India

<sup>2</sup>Student, Dept. of Computer Engineering, SVIT, Gujarat, India

\*\*\*

**Abstract** - Our project; Stock Market Analysis and Prediction, is a purely machine-learning-based model which is used to analyze the stock data of a provided firm through their previous historical stock data, and on the basis of that analysis, the future trend of the stock of that firm is predicted. Due to the current situation of the pandemic, more and more people are drawn towards investing in the stock market and so many new IPOs are coming up, hence we decided to build something that can help even novice people to properly invest their assets and money into some valuable stocks which can gain them some considerable amount of profit as every person aims for the maximum profit in the stock market.

**Key Words:** Stock Market, Stock Market Analysis, Stock Market Prediction, Machine Learning, Python, Graphs, NumPy, Scikit-Learn, TensorFlow, Streamlit, Matplotlib, FB Prophet, Plotly, Seaborn, CNN, RNN, LSTM, time-series

## 1. INTRODUCTION

Stock market; considering the vastness and the complexity it has, it's a very challenging task to comprehend especially for a novice person in this field. It holds a great risk and can lead to even greater loss if one invests in stock while not having any prior knowledge as the market is too volatile and uncertain to foresee or predict due to lots of fluctuations. The task of predicting the stocks is quite of an interesting endeavor in itself as it divides the academics and researchers into two sets; those who think that a mechanism can be devised to efficiently predict the stocks and those who think that the market is efficient and every time some new information comes up, the market absorbs it by correcting itself, thus there is no space for prediction and hence the first argument renders baseless.

Our project aims to carry out the study of the messy data that we get from the historical data set, train our model efficiently so that it can analyze the trends of the stock market accurately and can present a pictorial representation of the stock performance and can give the user an insight of the future trend of the stock and can provide the user with a favorable time period in which one can invest into the stocks and can reap more profit if not maximum.

Traditionally, almost all machine learning models use some or other observations as input features and they completely ignore the time-dimension of that data and thus, these traditional methods which are in use for a long time, they are not able to provide us with any significant time-based analysis which in return can help us predict the stocks in the near future.

To overcome this roadblock, a modern approach was adopted i.e., the Time-series forecasting model. This method, along with taking previous observations as input features, also considers the time-dimensions of the whole data set as the input feature and thus provides us with a much more accurate time-based analysis and prediction of the stock. The above-mentioned method is most widely adopted when we are dealing with the non-stationary data; the data for which its average mean value and its standard deviation value is not stationary or constant over time but it keeps fluctuating constantly and varies over time. Some classic examples of such time-series data are the varying temperature, the fluctuating stock market prices, the non-stationary rates of real-estate assets, etc.

Now, furthermore, the above-mentioned modern approach also has certain limitations when it comes to predicting the stock value behavior by using the simple linear time-series prediction method. Therefore, to tackle this challenge, an amalgamation of advantages of various methods can be used and we can use various algorithms. In this paper, we have used a hybrid of the time-series model and a combination of the CNN-LSTM model for forecasting the accurate trends of the stocks. Convolutional Neural Networks (CNN) basically works on extracting effective features from the given data set while Long Short-term Memory (LSTM) finds the interdependence of the data in time-series and automatically detects the best suitable approach for the given data. By combining the pros of these methods can effectively improve the accuracy of stock price forecasting. In order to check the efficacy and the accuracy of this model, this paper has used the quotidian transaction data of 3527 trading days of Apple Inc. from January 10, 2010, to January 20, 2020, in which the first 3027 trading days data are the training set and the last 500 trading days data are the test set. We have gathered this data from the open-source dataset provided by Yahoo Finance.

## 2. SCOPE

The Stock Market Analysis and Prediction model use a hybrid of different machine learning algorithms by combining the traditional and modern approaches in order to accurately predict the closing prices of the stock market. The model uses historical data from the data-set of any firm and time-dimensions as the input feature, it trains itself on the basis of the previous trends and fluctuations and undergoes training. The model then provides us with an in-depth analysis of the entire stock of the firm during that particular time period and then provides a forecast on the future trend of the stocks. We have used a tabular and a normal graphical representation to present the analysis and prediction of the stock. Additionally, we have also used OHLC candlestick graphs to provide a more comprehensible pictorial representation of the trend of the stocks to help our users understand the analysis and prediction properly.

## 3. OBJECTIVE

The objective behind building this stock market analysis and prediction model is:

- To accumulate a large enough data set of stocks of a firm and clean the data properly.
- To analyze the cleansed data and to train the model using a hybrid approach.
- To predict the future trend of the data using the data analyzed and to put it forward to the users.
- To present the overall analyzed data with its historical and future values to the users which are easy to comprehend.
- To predict the data with utmost accuracy in order to help the users to invest their capital properly in order to incur maximum profit.

## 2. LITERATURE REVIEW

There are quite a few approaches and methods that provide decent results for stock analysis and prediction. The proposed approaches have only focused on a few aspects of the model using traditional algorithms. A few of those citations giving a brief about the idea are:

**i. Title:** Machine Learning in Prediction of Stock Market Indicators based on Historical Data and Data from Twitter Sentiment Analysis

**Authors:** Alexander Porshney, Ilya Redkin, Alexey Shevchenko

**Description:**

- In this paper, the possibility to improve the accuracy of stock market indicators predictions by using data about the psychological states of Twitter users has been discussed.
- For the psychological state analysis, a lexicon-based method was adopted which helped in evaluating the eight basic sentiments in more than 755 million tweets.

**ii. Title:** Stock Market Forecasting using Machine Learning Algorithms

**Authors:** Haomiao Jiang, Tongda Zhang

**Description:**

- This paper talks about a new forecasting algorithm that makes use of the temporal correlation between the stock markets at the global level and different financial products to predict the next-day stock trend using the aid of SVM.

**iii. Title:** Stock Prices Prediction using Machine Learning and Deep Learning Techniques

**Author:** Aishwarya Singh

**Description:**

· This paper discusses predicting the future stock values. It starts off by discussing the use of simple machine learning techniques like linear and averaging regression and then proceeds further by discussing advanced approaches like auto ARIMA and LSTM.

**iv. Title:** Stock Market Prediction using Machine Learning Techniques

**Authors:** Edgar P., Torres P., Myriam Hernandez-Alvarez, Edgar A. Torres Hernandez, Sang Guun Yoo

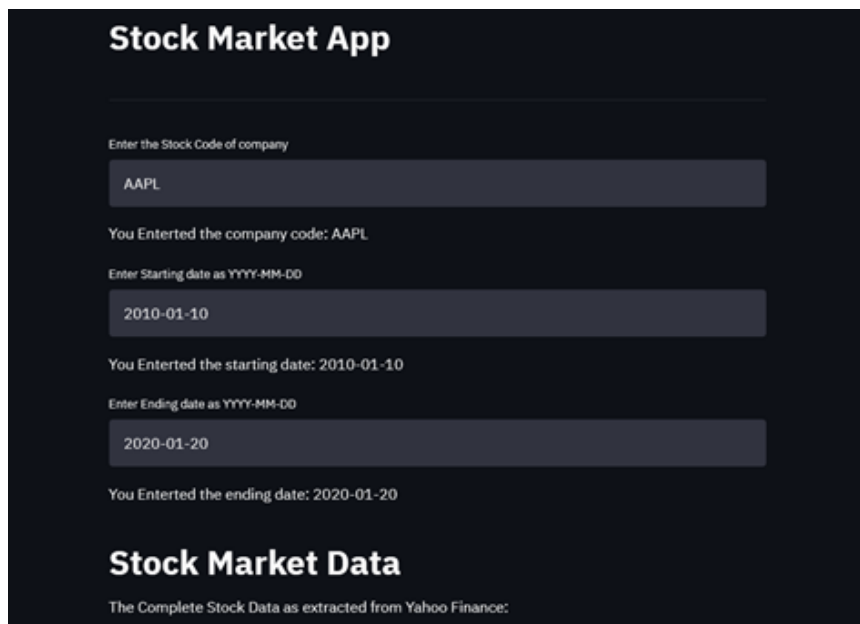
**Description:**

· This paper talks about using random trees and multilayer perceptron algorithm in order to predict the closing prices of the stock market.

**5. DESIGN AND METHODS**

i) The initial step is where the user enters the stock code of the company for which they want the model to analyze and predict the value of the stocks.

ii) Along with the stock code, the user has to enter the starting data and the ending data which will provide a time period to the model to analyze the data within that frame of time and to predict the values accordingly.



**Stock Market App**

Enter the Stock Code of company

AAPL

You Entered the company code: AAPL

Enter Starting date as YYYY-MM-DD

2010-01-10

You Entered the starting date: 2010-01-10

Enter Ending date as YYYY-MM-DD

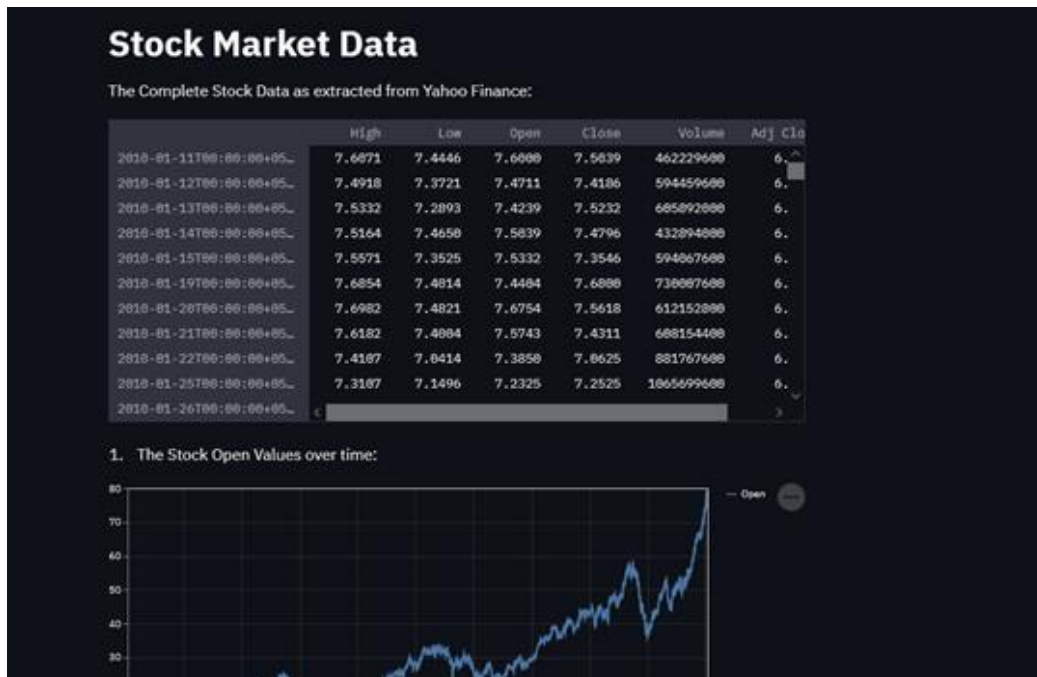
2020-01-20

You Entered the ending date: 2020-01-20

**Stock Market Data**

The Complete Stock Data as extracted from Yahoo Finance:

iii) The model then displays the complete stock data of the above-mentioned company in tabular representation which includes the high, low, open, close, volume, adjacent close values of that company. Here we have used the stock data of Apple Inc.



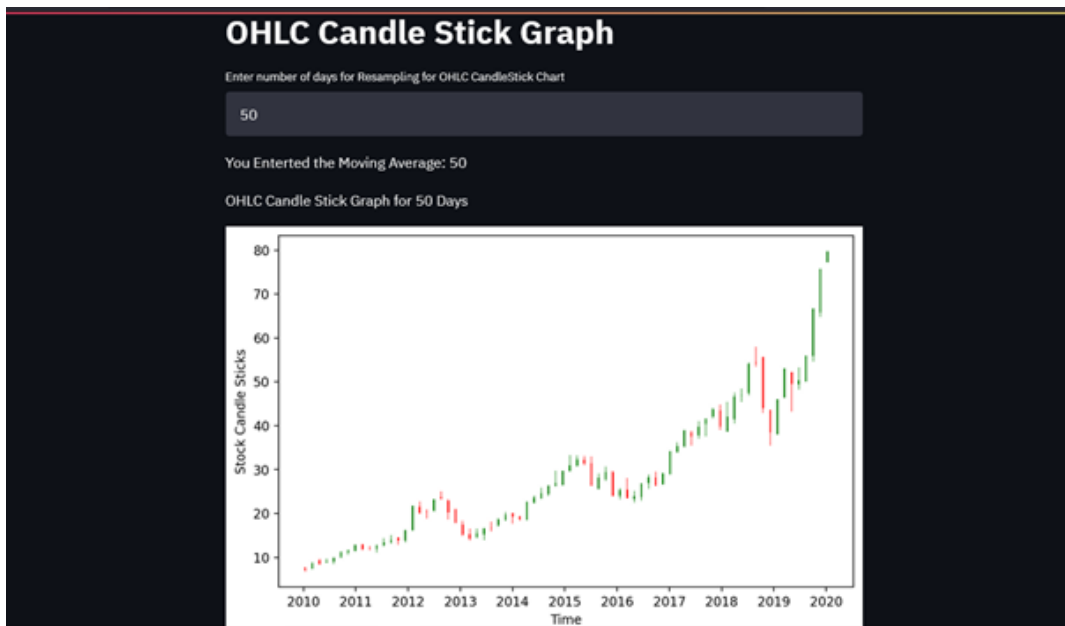
iv) Then the model displays the open and close values of the stock in the form of the graphical representation over the given time period.



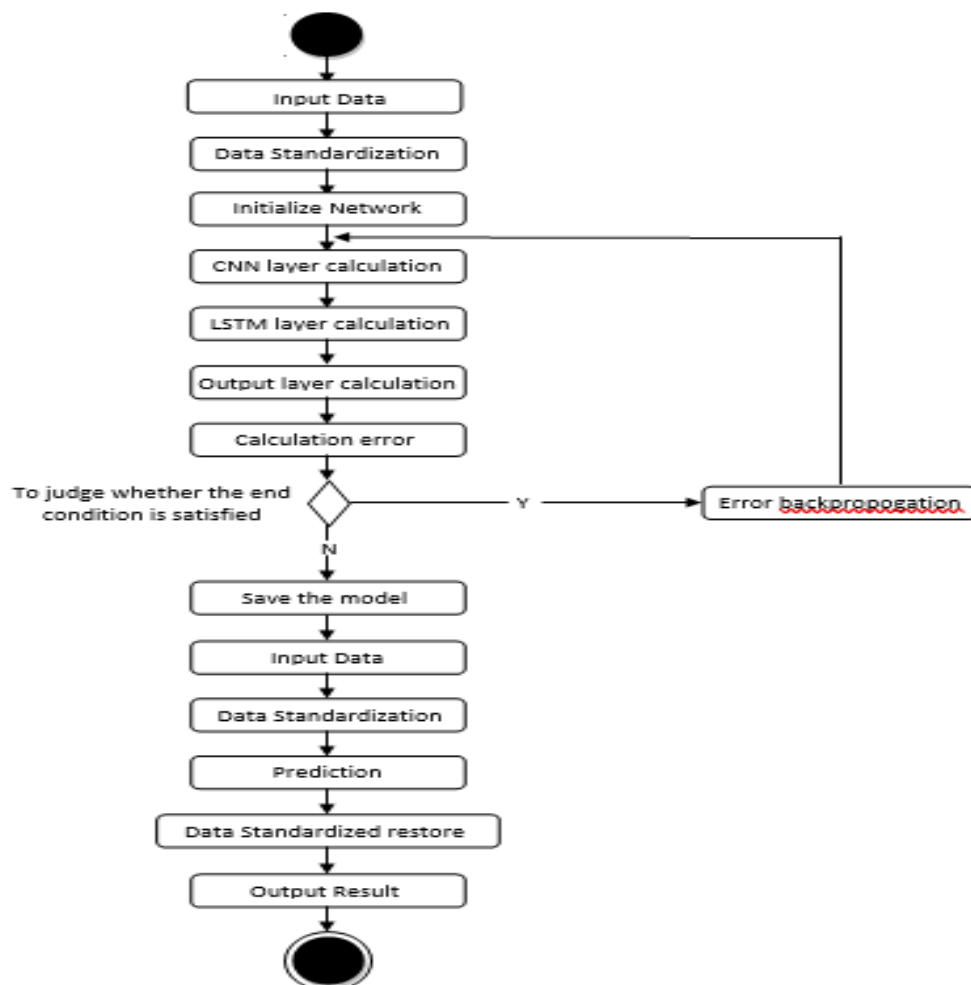
v) Then the user is asked to enter the moving averages which will predict the nearest and most accurate predicted value to that of the actual closing and opening values.



vi) On the basis of the given moving averages, an OHLC Candlestick Graph is also shown to the users to even get a more vivid and elaborate pictorial representation of the predicted values which the user can understand more easily.



### 6. FLOWCHART



## 7. CONCLUSION

Stock Market Analysis and Prediction is a Machine Learning based model which uses a hybrid approach by combining the traditional time-series and the modern LSTM, CNN, RNN, weighted and moving averages to increase the accuracy and predictability. The main function of the model is to provide the user with a thorough analysis of the stock of a specific firm which can provide a deep insight into the historical stock data of that firm and can predict the future trend of the stocks of that same firm. Due to the user-friendly interface and seamless user experience, the users can easily understand the presented data on their screen and can also easily fetch the required data of a stock of any firm.

## REFERENCES

- [1] Alexander Porshnev; Ilya Redkin; Alexey Shevchenko, "Machine Learning in Prediction of Stock Market Indicators Based on Historical Data and Data from Twitter Sentiment Analysis"
- [2] Haomiao Jiang; Tongda Zhang, "Stock Market Forecasting using Machine Learning Algorithms"
- [3] Aishwarya Singh, "Stock Prices Prediction Using Machine Learning and Deep Learning Techniques"
- [4] Edgar P; Torres P.; Myriam Hernandez-Alvarez; Edgar A. Torres Hernandez, "Stock Market Data Prediction Using Machine Learning Techniques"
- [5] Kunal Pahwa; Neha Agarwal, "Stock Market Analysis using Supervised Machine Learning"
- [6] Ashish Sharma; Dinesh Bhuriya; Upendra Singh, "Survey of stock market prediction using machine learning approach"
- [7] Ishita Parmar; Navanshu Agarwal; Sherish Saxena; Ridam Arora; Shikhin Gupta; Himanshu Dhiman; Lokesh Chouhan, "Stock Market Prediction Using Machine Learning"

## BIOGRAPHIES



Pranit Brahmbhatt

Student,

BE Computer Engineering,

SVIT Vasad,

Anand, Gujarat

- Web Developer, Application Developer, ML Enthusiast, Content Copywriter, UX Writer, Community Manager, and Tech Evangelist



Kushang Patel

Student,

BE Computer Engineering,

SVIT Vasad,

Anand, Gujarat

- ML Enthusiast, Budding Data Scientist